

ROSIE: Runtime Optimization of SPARQL Queries Using Incremental Evaluation

Lei Gai
lei.gai@pku.edu.cn

Wei Chen
pekingchenwei@pku.edu.cn

Tengjiao Wang
tjwang@pku.edu.cn

School of Electrical Engineering and Computer Science
Peking University, Beijing, P.R.China

ABSTRACT

Relational databases are widely adopted in RDF (Resource Description Framework) data management. For efficient SPARQL query evaluation, the legacy query optimizer needs reconsiderations. One vital problem is how to tackle the suboptimal query plan caused by error-prone cardinality estimation. Consider the schema-free nature of RDF data and the JOIN-intensive characteristic of SPARQL query, determine an optimal execution order before the query actually evaluated is costly or even infeasible, especially for complex queries on large-scale data. In this paper, we propose ROSIE, a Runtime Optimization framework that iteratively re-optimize SPARQL query plan according to the actual cardinality derived from Incremental partial query Evaluation. By introducing an approach for heuristic-based plan generation, as well as a mechanism to detect cardinality estimation error at runtime, ROSIE relieves the problem of biased cardinality propagation in an efficient way, and thus is more resilient to complex query evaluation. Extensive experiments on real and benchmark data show that compared to the state-of-the-arts, ROSIE consistently outperformed on complex queries by orders of magnitude.

CCS Concepts

•Database Management → Systems; Query processing;

Keywords

SPARQL query optimization, cardinality estimation, runtime optimization

1. INTRODUCTION

The RDF¹ data model and SPARQL² query language are W3C (World Wide Web Consortium) recommended standards for representing and accessing the ever-increasing linked

resources, e.g., DBpedia [1], Freebase [2], UniProt [3], etc. In practice, RDF data are physically managed by infrastructures, which can be dedicated RDF stores developed from scratch, e.g., [28,35,37], or using legacy relational databases (RDB) as back-end, e.g., [4,7,8,14]. RDB-backed systems can benefit greatly from 40+ years of advances in database domain, and have been an active research topic in both research literatures and industries. However, in the context of RDF/SPARQL, query optimization strategies directly borrowed from relational optimizer are not universally applicable [18]. This is especially true for complex queries on large-scale data. In general, there are two noteworthy problems that need to be carefully reconsidered.

First, due to the schema-free nature of RDF, cardinality estimation is error-prone, and sometimes even biased by orders of magnitude. In a traditional query optimizer, cardinality estimation is the dominating component that affects optimal query plan selection [25]. It is canonically based on the containment and independence assumption. To tackle correlations in data, the most widely adopted and cost-effective method is by maintaining pre-build statistics as multi-dimensional histograms [19]. For RDF data, it is commonly represented as a collection of *triples* in the form of three elements, $\langle \text{Subject}, \text{Predicate}, \text{Object} \rangle$, (or $\langle \text{S}, \text{P}, \text{O} \rangle$ for brevity). Even though comprehensive statistics of all joinable combinations can exactly capture correlations in such data, but they tend to be exponentially huge, as each triple may has correlations with others. Existing researches proposed RDF-specific statistical synopses, such as permutation indexes [35], possible join path [28] or *Characteristic Sets* [27], in order to give better estimates than always assuming independence. But they are still insufficient for more complex queries. Besides, managing and accessing these statistics needs a redesign of the underlying RDF storage. This is not well suited for RDB-backed systems.

Second, consider the SPARQL's JOIN-intensive characteristic, the propagation of cardinality estimate errors has a more profound effect on query performance. A SPARQL query is typically specified as a combination of *triple patterns*³ (TP) in a fine-grained manner. The query plan (a.k.a., execution order) consists of significantly more JOINS caused by each single access of a TP. Earlier work has shown that errors in cardinality estimate may increase exponentially with the number of JOINS [21]. Although carefully cultivated storage, e.g., property tables [11,36] or vertically partitioned columns [4], can alleviate this problem by reducing the number of JOINS, they did not solve the fundamental problem

¹<https://www.w3.org/RDF/>

²<http://www.w3.org/TR/sparql11-query>

³<https://www.w3.org/TR/sparql11-query/#QSynTriples>

of cardinality estimate error. Furthermore, in real-world queries, syntactic constructs, such as OPTIONAL, UNION and FILTER etc, are extensively used [31]. These complicated the problem.

To tackle these problems, runtime optimization approaches [6, 9, 12, 15, 26] are promising. This kind of approaches incrementally materialize the intermediate results during query execution time, gather the accurate cardinalities and eliminate errors at given steps, then use these feedbacks to dynamically adjust the query plan. In this case, cost-efficient heuristics-based methods [16, 27, 33] are feasible for query plan generation. Such plan is not guaranteed to be optimal, but can be dynamically optimized at runtime. Nonetheless, consider the pervasive existence of interactions between JOIN predicates in a SPARQL query, directly applying above methods may results in a materialization at each query execution step. This can greatly down-grades the SPARQL query performance as materialization is costly. Therefore, two challenges arise: (i) how to generate an optimized TP execution order that needs as fewer materialization steps as possible; (ii) How to determine at which step an incremental materialization is needed during runtime optimization.

In this paper, we propose ROSIE, a framework for Runtime Optimization of SPARQL query using Incremental Evaluation. The basic idea of ROSIE is based on that, as an optimal execution order is hard to be pre-determined in a cost-effective way, runtime evaluation of partial query at a given step can eliminate the cardinality estimation error that may propagate to the following steps, and materialized intermediate results can be used to guide the optimal selection of the remaining execution order. It starts with an initial TP execution order, named *Candidate Sequence (CS)*, then sequentially determines the cardinality estimation error at each step in *CS*. Incremental evaluation of partial query is triggered once ROSIE detects that current step may lead to a suboptimal execution. ROSIE follows a RDB-backed infrastructure. This guarantees the efficiency of *CS* evaluation and intermediate results materialization. The following techniques tackle the above challenges brought by adopting runtime optimization.

For the first challenge, we propose a heuristic-based method for *CS* generating. Different from existing works that mainly focus on TP JOINS, we introduce a query graph, named as *Query Relation Graph (QRG)*, to model the query structure of more versatile semantics in SPARQL that includes UNION, FILTER and OPTIONAL, etc. To generate a *CS* that optimally guides the query execution, we present a greedy algorithm to traversal QRG, which synthesizes the semantical equivalence of SPARQL operators and preliminary statistics of RDF data. This is detailed in *Section 5*.

For the second challenge, we devise a mechanism for possible suboptimal steps detection. It utilizes the propagation of cardinality estimation errors in a *CS*. We argue that, for an optimal *CS*, the estimation error bound is propagated in a non-incremental fashion. Consider that the error estimate needs not to be as accurate as the cardinality estimate, the general statistics maintained by a RDB are useful in determining a feasible error bound. This makes ROSIE more practical as it does not need to revise the implementation of the underlying RDB. This is detailed in *Section 6*.

To summarize, our contributions in this paper are three-fold: (i) We introduce a radical new runtime optimization framework, named ROISE. It uses incremental partial query

evaluation to dynamically optimize the query execution, and is characterized by supporting efficient evaluation of more expressive SPARQL queries; (ii) To boost the performance of ROSIE, we propose a heuristics-based approach for generating an optimized TP execution order, and devise a mechanism that leverage error bounds estimate for possible suboptimal steps detection; (iii) Through extensive experiments on benchmark and real datasets, we show the correctness and efficiency of ROSIE. For complex queries, ROSIE yields performances that are 1.5-20 times better than the state-of-the-arts.

The rest of the paper is organized as follows: *Section 2* reviews the most related works. *Section 3* introduces the principal concepts used. *Section 4* gives an overview of ROSIE. *Section 5* details the heuristic-based method for *CS* planning, and the mechanism for determining the error bound propagation is introduced in *Section 6*. *Section 7* shows the experimental results, and *Section 8* concluded.

2. RELATED WORK

Cardinality estimation is still an active topic, as it tends to be more error-prone among all the costs involved in query optimization [10]. Existing literatures can be roughly classified as *static* approaches or *runtime* approaches.

Static approaches rely on off-line maintained statistics, such as histograms (e.g., [19]) or samples (e.g., [24]), etc, to determine a query plan that is unalterable throughout the query execution. In the context of RDF/SPARQL, maintaining all possible statistics is unrealistic. Statistics of permutation indexes [35], or even more complicated statistics, such as *Characteristic Sets*, [27] are proposed. They only delay the step where errors propagated, thus are less effective for complicated queries with more JOINS. Heuristics-based approaches, e.g., using greedy strategy [33] or query structure [34], are effective in pruning the plan search space, and are extensively adopted as subsidiary mechanisms [16, 28, 37]. But they are not guaranteed to be optimal.

Other than attempting to pre-determine an optimal query plan, *Runtime* approaches [6, 9, 12, 15, 22, 26] optimize query plan at runtime when the actual cardinality turns available. There are few works of *Runtime* optimization for RDF/SPARQL. The most relevant works in literatures are [6], [22] and [26]. [6] materialized intermediate results at each step, which seems to be an overkill. [22] proposed a similar framework as ROSIE, but is based on [6] and focused on dealing with skews in a distributed scenario. [26] is the most similar to ROSIE, as they all devise mechanism for detecting error propagation, which are used to selectively choose plan fragments to be executed. But the heuristics adopted in [26] generally work well in RDBs, they need to be reconsidered for SPARQL queries.

3. PRELIMINARIES

3.1 RDF and SPARQL

In RDF data model, a real world *resource* is identified by an unique URI (Uniform Resource Identifier) string, and each of its properties or linkages with other *resources* is depicted as a triple t_i of three components, $t_i = \langle S, P, O \rangle$. A collection of such triples forms an *RDF dataset* $\mathcal{D} = \bigcup_{i=1}^{|\mathcal{D}|} t_i$.

Like SQL for relational database, SPARQL is the de facto query language for RDF. Basically, a SPARQL query con-

SELECT ?u1 ?u2 ?p1 WHERE {	Cardinality
?u1 type User .	T_1 30
?u2 type User .	T_2 30
?u1 creator_of ?p1 .	T_3 200
?p1 type Post .	T_4 70
?p1 hashtag "Database" .	T_5 15
?p1 content ?pc .	T_6 500
{ ?u1 knows ?u2 . }	T_7 120
UNION	
{ ?u1 knows ?u3 .	T_8 120
?u3 knows ?u2 . }	T_9 120
OPTIONAL	
{ ?u2 like ?p1 .	T_{10} 350
?u1 reply_of ?p1 . }	T_{11} 80
FILTER (?pc < 39) }	C 20

Figure 1: An example SPARQL query Q_e .

$\mathbb{T} = \{T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9, T_{10}, T_{11}\}$
 $\mathbb{O} = \{\text{AND}_1, \text{AND}_2, \text{AND}_3, \text{AND}_4, \text{OR}, \text{OPT}, \text{FILTER}\}$
 $\mathbb{V} = \{ ?u1 \xrightarrow{S_1} (T_1, T_3, T_7, T_8, T_{11}),$
 $\quad ?u2 \xrightarrow{S_2} (T_2, T_{10}), \quad ?u2 \xrightarrow{O} (T_7, T_9),$
 $\quad ?u3 \xrightarrow{S_3} (T_9), \quad ?u3 \xrightarrow{O} (T_8),$
 $\quad ?p1 \xrightarrow{S_4} (T_4, T_5, T_6), \quad ?p1 \xrightarrow{O} (T_3, T_{10}, T_{11}),$
 $\quad ?pc \xrightarrow{S_5} (T_6) \}$
 $\mathbb{S} = \{ \text{AND}_1 \rightarrow (\text{OR}, \text{AND}_2, \text{FILTER}, \text{OPT}),$
 $\quad \text{AND}_2 \rightarrow (T_1, T_2, T_3, T_4, T_5, T_6),$
 $\quad \text{OR} \rightarrow (T_7, \text{AND}_3),$
 $\quad \text{AND}_3 \rightarrow (T_8, T_9),$
 $\quad \text{AND}_4 \rightarrow (T_{10}, T_{11}),$
 $\quad \text{OPT} \rightarrow (\text{AND}_4), \quad \text{FILTER} \rightarrow (C) \}$

Figure 2: Formalization of $Q_e = \langle \mathbb{T}, \mathbb{O}, \mathbb{V}, \mathbb{S} \rangle$.

sists of a query form (i.e. SELECT, ASK, CONSTRUCT or DESCRIBE), a *Basic Graph Pattern* (BGP) specified in a WHERE clause, and possibly solution modifiers (e.g. DISTINCT). The building block of BGPs are TPs. Each TP is expressed in a triple form, with a least one of S, O, P elements replaced by unbounded *variables*⁴. A BGP can be recursively defined as a TP, or a finite set of BGPs connected by SPARQL operators (AND, OR, OPT, and FILTER) to build expressive queries⁵. A TP JOIN is a JOIN (correspondence to AND operator, analogous to equi-JOIN) between two TPs on a variable (or a join predicate in relational database literatures). A SPARQL query Q can be formalized as a quadruple, $Q = \langle \mathbb{T}, \mathbb{O}, \mathbb{V}, \mathbb{S} \rangle$, where,

- \mathbb{T} is the finite set of TPs in Q .
- \mathbb{O} is the set of SPARQL operators, For $\odot \in \mathbb{O}$, \odot can be one of AND, OR, OPT or FILTER.
- \mathbb{V} is the *variable correlations*, expressed as groups of TPs in \mathbb{T} that share common variables at given positions (S, O or P).
- \mathbb{S} is the *query semantics*, which states how TPs in \mathbb{T} are connected by operators in \mathbb{O} .

EXAMPLE 1. Figure 1 shows an example of a SELECT query Q_e . In its WHERE clause, 11 TPs and a FILTER constraint is defined⁶. The formalization of Q_e is shown as Figure 2.

⁴<https://www.w3.org/TR/sparql11-query/#QSynVariables>

⁵Specifically, AND connects two conjunctive BGPs, OR for BGPs connected by UNION keywords, OPT for OPTIONAL keywords, and FILTER for FILTER conditions.

⁶Noteworthy that in order to make this example easier to read, we replaced the URIs with more readable names in Q_e .

3.2 Candidate Sequence (CS)

Traditionally, a SPARQL query Q is evaluated in a TP-by-TP manner. The optimization goal is focused on determining an optimal operator tree of TPs that minimize the intermediate results size during query evaluation. In this paper, we formally define an representation of the operator tree, named as a *Candidate Sequence* (CS), as follows :

DEFINITION 1 (CANDIDATE SEQUENCE). Given a $Q = \langle \mathbb{T}, \mathbb{O}, \mathbb{V}, \mathbb{S} \rangle$, a *Candidate Sequence* CS can be recursively defined as:

- (T) is a CS, where $T \in \mathbb{T}$.
- If CS_m, CS_n are CS, for $\odot \in \mathbb{O}$, $(CS_m \odot CS_n)$ is also a CS.

The concept of CS are analogous to the query plan used in the relational *cost-based query optimization* (CBO). A partial CS can be regarded as a *step* in evaluation. As CS is expressed in the form of SPARQL algebra, like the relational equivalence query plans in CBO, a Q may have many *equivalent* (denoted by \equiv) CSs [29]. Let T_1, T_2 and T_3 be TPs and C is a constraint, the following *Properties* hold.

PROPERTY 1. (*Exchangeable Region*) The position of a TP in an AND-only or an OR-only CS is exchangeable.

1. $((T_1 \text{ AND } T_2) \text{ AND } T_3) \equiv ((T_1 \text{ AND } T_3) \text{ AND } T_2)$.
2. $((T_1 \text{ OR } T_2) \text{ OR } T_3) \equiv ((T_1 \text{ OR } T_3) \text{ OR } T_2)$.

Property 1 states that TPs connected by AND or OR can freely exchange their positions in a CS. It stands as both AND and OR are commutative and associative.

PROPERTY 2. (*Distributable Region*)

1. $(T_1 \text{ AND } (T_2 \text{ OR } T_3)) \equiv ((T_1 \text{ AND } T_3) \text{ OR } (T_1 \text{ AND } T_2))$.
2. $(T_1 \text{ OPT } (T_2 \text{ OR } T_3)) \equiv ((T_1 \text{ OPT } T_3) \text{ OR } (T_1 \text{ OPT } T_2))$.
3. $((T_1 \text{ OR } T_2) \text{ OPT } T_3) \equiv ((T_1 \text{ OPT } T_3) \text{ OR } (T_2 \text{ OPT } T_3))$.
4. $((T_1 \text{ OR } T_2) \text{ FILTER } C) \equiv ((T_1 \text{ FILTER } C) \text{ OR } (T_2 \text{ FILTER } C))$.

Property 2 states that a latter TP can be distributed to its proceeding steps in CS. Its correctness can be proven according to the SPARQL algebra equivalence stated in [29]. Property 1 and 2 are the foundations of equivalent CS generation used in this paper.

Recap that in CBO, cardinality estimation is performed at each step of a query plan to derive a cardinality of the contemporary result set (a.k.a. a cost). Given an RDF dataset \mathcal{D} , and a SPARQL query $Q = \langle \mathbb{T}, \mathbb{O}, \mathbb{V}, \mathbb{S} \rangle$, we detail the problem of cardinality estimation in a CS from two aspects: the cardinality of TPs, and the cardinality of TP JOINS.

The cardinality of a TP $T_i = \langle S, P, O \rangle \in \mathbb{T}$ is determined as $|T_i| = p(T_i) \times |\mathcal{D}|$, where $p(T_i)$ is the selectivity of T_i , and can be explained as the distribution of triples in \mathcal{D} that match T_i . Clearly, $p(T_i) \in [0, 1]$. Under the independence assumption (i.e. S, P and O are mutually independent), $p(T_i)$ can be estimated by the formula $p(T_i) = p(S) \times p(P) \times p(O)$, where $p(\bullet) = 1.0$ if element $\bullet \in \{S, P, O\}$ is unbounded, and for bounded object, $p(\bullet)$ is approximated by means of classical single-value histograms in a RDB [33]. In fact, independence assumption rarely holds in \mathcal{D} .

The cardinality of a TP JOIN can be varied between 0 and the Cartesian product size. The join selectivity estimation is a further investigated problem for RDB with a well-defined schema. It general, the cardinality of AND operator can be estimated as :

$$|T_i \text{ AND } T_j| = p(T_i \text{ AND } T_j) \times |T_i| \times |T_j| \quad (1)$$

where $p(T_i \text{ AND } T_j)$ denotes the selectivity of AND. Under the containment assumption, $p(T_i \text{ AND } T_j) = \frac{1}{\max(p(T_i), p(T_j))}$ [10]. This works well for foreign-key join. Consider the schema-free nature of \mathcal{D} , such estimate is error-prone.

An incorrect single step in cardinality estimation potentially screws up whole execution plan. As maintaining comprehensive statistical informations in a RDB is costly and impractical, cardinality estimation errors are unavoidable. We define the cardinality estimation error as:

$$\varepsilon_{\odot}(T_i, T_j) = \frac{|T_i \odot T_j|_P}{|T_i \odot T_j|_E} \quad (2)$$

, where $|\cdot|_P$ denotes the real cardinality, and $|\cdot|_E$ denotes the estimated cardinality. $\varepsilon < 1$ means an over-estimate, and $\varepsilon > 1$ means an under-estimate. We further discuss the error estimation in Section 6.2.

EXAMPLE 2. Consider Q_e in Figure 1. We can assign a $CS = ((T_1 \text{ AND } T_3) \text{ AND } T_6)$ for partial query evaluation. There $CS_{Step1} = (T_1 \text{ AND } T_3)$, and $CS_{Step2} = (CS_{Step1} \text{ AND } T_6)$. Adopting traditional cardinality estimates method discussed above, we get $|CS_{Step1}| = 30$, $|CS_{Step2}| = 30$. Consider the semantic correlations exists in \mathcal{D} , that is each Post is created by a User, and each Post has a content, the actual cardinalities are $|CS_{Step1}| = 200$, $|CS_{Step2}| = 500$. Thus, $\varepsilon_{AND}(T_1, T_3) = \frac{200}{30} \approx 6.7$, $\varepsilon_{AND}(CS_{Step1}, T_6) = \frac{500}{30} \approx 16.7$. This states that current estimates are vastly under-estimated.

4. OVERVIEW OF ROSIE

The general idea of ROSIE boils down to determining the optimal CS dynamically during query evaluation. ROSIE is based on the rationale of pushing as much hard and time consuming operations into the RDB as possible. This can profit most from the existing highly optimized techniques, while using runtime optimization technique to alleviate the problems caused by RDF/SPARQL-specific features. The main algorithm is shown in Algorithm 1. It mainly consists of two phases.

Algorithm 1: ROSIE algorithm.

```

input :  $Q$ , and  $\mathcal{D}$  over a RDB.
output:  $R$ , evaluation results of  $Q$  on  $\mathcal{D}$ .
begin
1   $Q$  parsing, and QRG  $G_Q$  construction; //Section 5.1;
2   $CS \leftarrow CSConstruction(G_Q)$ ; //Section 5.2;
3   $T_1, T_2 \leftarrow$  Get the first 2 TPs from  $CS$ ;
4   $CS_{sub} \leftarrow \{T_1\}; R_s \leftarrow \{\cdot\}; \widehat{G_Q} \leftarrow G_Q$ ; //Initialization;
5  while having TPs not traversed in  $G_Q$  do
6     $\varepsilon \leftarrow CardErrEsti(CS_{sub}, T_i, T_{i+1})$ ; //Section 6.2;
7    if error bound exceeded then
8       $CS_{sub} \mapsto SQL$ ; //Convert  $CS_{sub}$  to SQL;
9       $R_i \leftarrow SQL$  evaluation in RDB;
10      $CS_{sub} \leftarrow \{T_i\}; R_s \leftarrow R_i; \widehat{G_Q} \leftarrow$  update  $\widehat{G_Q}$ ;
11      $CS \leftarrow CSConstruction(\widehat{G_Q})$ ; //Section 5.2;
12   else
13      $CS_{sub} \leftarrow \odot, TP_i$ ;
14      $T_i \leftarrow$  Get next TP from  $CS$ ;
15 if  $CS_{sub}$  is not null then
16    $CS_{sub} \mapsto SQL$ ;  $R \leftarrow SQL$  evaluation;
17 else
18    $R \leftarrow R_s$ ;

```

The first phase analyzes the structure of a query and initializes a CS . In essence, after a SPARQL query Q is submitted and parsed, a *Query Relation Graph* (QRG) that models

\mathbb{V} and \mathbb{S} in Q is constructed (Line 1, details in Section 5.1). Then an initial CS is constructed by traversing the QRG following to a set of heuristics in a greedy manner (Line 2, details in Section 5.2).

The second phase performs runtime optimization during query evaluation. It is achieved by iteratively determining an cardinality estimate error bound at each step of CS . Once the augmentation of current TP tends to cause cardinality estimate unacceptably biased, the equivalent SQL of the partial CS that contains all ancestor TPs is generated and submitted to the underlying RDB for execution (Line 6-11, details in Section 6.2). The materialized results of the partial CS evaluation, as well as data statistics, are managed by the underlying RDB, and are used in the next iteration to plan CS for the remaining part of query at runtime (Line 10-11). Otherwise, the algorithm just proceed with the next TP in CS (Line 13-14). This phase iterates until all TPs have been processed, and returns the final results (Line 18-18).

In Algorithm 1, to minimize the materialization costs, ROSIE focus on the optimization of two key operations: the CS planning (Line 2 and 11) and the cardinality estimate error determination (Line 6). We details them in Section 5 and 6 respectively.

5. CANDIDATE SEQUENCE PLANNING

In this section, we introduce a method to generate an optimized CS . It models the query structures into a graph structure, named as *Query Relation Graph* (QRG). We explain the formation of QRG in Section 5.1, and devise a novel linear-time heuristics-based greedy algorithm for CS planning in Section 5.2.

5.1 Query Relation Graph (QRG)

As discussed in Section 3.1, a SPARQL query Q have its *variable correlations* \mathbb{V} and *query semantics* \mathbb{S} . To well model Q , we introduce the concept of Query Relation Graph (QRG). It can be formally defined as :

DEFINITION 2 (QUERY RELATION GRAPH, QRG). *The QRG of a SPARQL query Q is a directed, vertex-weighted, edge-labeled graph, denoted as $G_Q = (V, E, f_w, f_l)$, where:*

- *Vertex set $V = V_I \cup V_{II} \cup V_{III}$. $v_{\odot} \in V_I$ denotes an operator, $v_t \in V_{II}$ denotes a triple pattern, and $v_v \in V_{III}$ denotes an unbounded variable in Q .*
- *Edge set $E = E_I \cup E_{II}$. $e \in E_I$ is a direct edge from $v_t \in V_{II}$ to $v_{\odot} \in V_I$, and $e \in E_{II}$ is a direct edge from $v_v \in V_{III}$ to $v_t \in V_{II}$, labeled by the position of the variable v_v in triple pattern v_t , i.e. S, P , or O .*
- *Weight function $f_w : w \rightarrow V_{II}$ assigns each vertex $v \in V_{II}$ a weight $w(v)$.*
- *Label function $f_l : l \rightarrow E_{II}$ assigns each edge $e \in E_{II}$ a label $l(e)$, $l(e) \in \{S, O, P\}$.*

Intuitively, in a QRG, the subgraph $G_{\mathbb{V}} = (V_{II} \cup V_{III}, E_I, f_l)$ models the *variable correlations* \mathbb{V} , and the subgraph $G_{\mathbb{S}} = (V_{II} \cup V_{III}, E_{II})$ models the *query semantics* \mathbb{S} . The data statistics is introduced by the weight function f_w .

The construction of QRG is quite straightforward. Firstly, $G_{\mathbb{V}}$ can be structured following the operator tree derived from the standard SPARQL query parser (e.g. *Rasqal RDF*

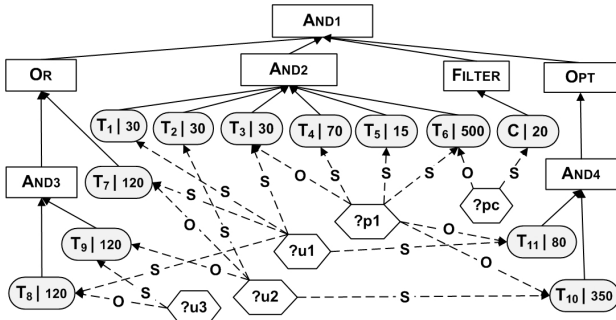


Figure 3: Query Relation Graph for Q in Figure 1.

Query Library⁷), only adding direction of edges from child vertex to its parent vertex in a bottom-up manner. Then, the estimated cardinality of each triple pattern $T \in \mathbb{T}$ is assigned as a weight to the corresponding vertex $v_t \in V_{II}$. Following, for each unbound variable in Q , a vertex $v_v \in V_{III}$ is added to G_Q . A directed edge from v_v to $v_t \in V_{II}$ which represents the TP $T \in \mathbb{T}$ that contains this variable is established, and labeled with the component position of this variable (S, O or P) in T .

EXAMPLE 3. Figure 3 shows a QRG of Q_e in Figure 1. It consists of three types of vertices, which are depicted by different kinds of symbols.

5.2 Heuristic-Based \mathcal{CS} Planning

Analogous to the query plan selection in CBO, the aim of optimal \mathcal{CS} planning is to select the one with the minimal cost from the search space of all equivalence \mathcal{CS} s. This is a classic problem in query optimization, and can be reduced to the problem of finding a minimal weight tree that covers all $v_t \in V_{II}$ in QRG. It is NP-hard as it can be deduced from the well-known TSP problem. Exhaustive exploration is impractical, as Q may contains tens or even hundreds of TPs. Besides, the techniques borrowed from traditional SQL query optimization may not works well due to lack of schema or integrity constraints.

Nonetheless, consider that \mathcal{CS} may be dynamically optimized at runtime, we advocate the use of heuristics-based approach to give a reasonable \mathcal{CS} that can be used to guide further optimizations. To better represent the *query semantics* \mathbb{S} which includes operators like OR, OPT and FILTER besides AND, we introduce three concepts, namely *Region*, *Ancestor* and *Available Variable*.

Region: A *Region* is comprised of all TPs that connects to the same operator $v_{\odot} \in V_I$. The relative position of a *Region* is decided by the position of $v_{\odot} \in V_I$ in G_V . More formally, A *Region* $R(\odot)$ of operator \odot is defined as follows:

$$R(\odot) = \{v_t | \exists e(v_t \rightarrow v_{\odot}), e \in E_I, v_{\odot} \in V_I, v_t \in V_{II}\}$$

Ancestor and LCA: For $v_v \in V_{III}$, the set of its *Ancestor*, $A(v_v)$, contains all the reachable (denoted as \rightsquigarrow) operators $v_{\odot} \in V_I$, and its *Least Common Ancestor* (LCA), $LCA(v_v)$, is an operator $v_{\odot} \in V_I$ that is reachable to all $v_{\odot} \in A(v_v)$ with the shortest distance.

$$A(v_v) = \{v_{\odot} | \exists v_v \rightsquigarrow v_{\odot}, v_{\odot} \in V_I, v_v \in V_{III}\};$$

$$LCA(v_v) = \operatorname{argmin}_{\sum d(v_{\odot} \rightsquigarrow \hat{v}_{\odot})} (\hat{v}_{\odot}), v_{\odot} \in A(v_v), \hat{v}_{\odot} \in V_I.$$

⁷<https://github.com/dajobe/rasqal>

where $d(v_{\odot} \rightsquigarrow \hat{v}_{\odot})$ is the distance of the shortest path from v_{\odot} to \hat{v}_{\odot} in G_Q .

Available variable: For $v_v \in V_{III}$, it is called as an *available variable* in a *Region* $R(\odot)$, if all its connected $v_t \in V_{II}$ in $R(\odot)$ have been selected as candidates in \mathcal{CS} .

EXAMPLE 4. For G_Q shown as Figure 3, T1-T6 are all connected to the AND2 operator, they are in a *Region*, $R(\text{AND2})$. Furthermore, according to Property 1, $R(\text{AND2})$ is an *Exchangeable Region*. For vertex $?u1 \in V_{III}$, $A(?u1) = \{\text{AND1}, \text{AND2}, \text{OR}, \text{OPT}, \text{AND3}, \text{AND4}\}$, and $LCA(?u1) = \text{AND1}$. Consider $\widetilde{\mathcal{CS}} = (((T_5 \text{ AND } T_4) \text{ AND } (T_6 \text{ FILTER } C)) \text{ AND } T_3)$, $?p1$ is available in $R(\text{AND2})$, as all its connected $v_t \in V_{II}$ (i.e., T_4, T_5, T_6, C) have been arranged in $\widetilde{\mathcal{CS}}$. While for $\widetilde{\mathcal{CS}}$, $?u1$ is not available in $R(\text{AND2})$ until T_1 is arranged.

Using the concepts introduced above, we propose the heuristics adopted in \mathcal{CS} planning. The rational behind heuristic-based method is due to that RDF data always have certain fixed patterns among triples that fits the structure of Q . Using these particular features can produce promising results, as stated in [33, 34]. Based on this, we propose the following heuristics:

- **H1:** The vertices $v \in V_{II}$ is ordered by the labels of their incoming edges as: $\emptyset \prec P \prec S \prec O \prec S, P \prec P, O \prec S, O \prec S, P, O$, where \emptyset means no incoming edges. Additionally, for $v \in V_{II}$ with the same types of incoming edges, the less weighted vertex is preferential. This heuristic is a generalization of H1 defined by [34] on condition that the basic summary data is available.
- **H2:** $v_t \in V_{II}$ in the same *Exchangeable Region* and connected to the same $v_v \in V_{III}$ using the same label can be arranged in a successive manner following H1. If v_v has been *Available* in a *Region*, then $v_t \in V_{II}$ in other *Regions* and is connected with v_v can be chosen with priority. This heuristic favours the star-shaped subqueries.
- **H3:** Evaluates FILTER at the earliest possible point. A $v_v \in V_{III}$ that is related to more FILTERS has a higher selectivity.
- **H4:** For an $R(\text{OPT})$, all TPs can not be planned unless all of the $v_v \in V_{III}$ that these TPs connected with are *available*. This is because that OPT corresponds to LEFT OUTER JOIN, which is not associative or distributive over other operators.

Adopting these heuristics, we propose a greedy \mathcal{SC} generation algorithm, given as *Algorithm 2*. The algorithm can be divided into the *Arrange* phase and the *Generation* phase.

In the *Arrange* phase, it begins with picking the most selective TP $T^{(1)}$ using H1, and $T^{(1)}$ is assigned to T^c as the seed TP to start with (Line 1). Then, from the *Region* $R(\odot)$ that contains T^c , all TPs that are not arranged in \mathcal{CS} are selected in the candidate set, V^c (Line 3). Meanwhile, the *available variables*, i.e., $v_v \in V_{III}$ that connected to the arranged TPs, is maintained in a set V^v (Line 9). If V^c is an empty set, which states that all TPs in $R(\odot)$ have been arranged, the algorithm explores all available $v_v \in V_{III}$, traverses to the *Regions* that contains TPs which are connected with these *available variables*. Then all TPs contained in these *Regions* are held in V^c as candidates (Line 4-8). After determining V^c and V^v for current T^c , the connected FILTER operators are checked and arranged if exists using H3 (Line 10-11). After that, all connected OPT operators are also checked. The arrangement of OPT related TPs is

Algorithm 2: Greedy algorithm for \mathcal{CS} generation.

```

input :  $G_Q$ , the QRG of query  $Q$ .
output:  $\mathcal{CS}$ , a candidate sequence.
begin
1   $T^{(1)} \leftarrow$  chosen according to H1;  $V_{II}^a \leftarrow \{T^{(1)}\}$ ;  $T^c \leftarrow T^{(1)}$ ;
2  while  $V_{II}^a \subset V_{II}$  do
3     $V^c \leftarrow \{(v^c | T^c \subset R(\odot) \wedge v^c \in R(\odot)) \wedge (v^c \notin V_{II}^a)\}$ ;
4    if  $V^c$  is  $\emptyset$  then
5       $V_{III}^a \leftarrow \{v | v \rightarrow v_t, v \in V_{III}, v_t \in V_{II}^a\}$ ;
6      for  $\forall v \in V_{III}^a$  do
7         $V_{II}^c \leftarrow \{v^c | v \rightarrow v^c, v \in V_{III}^a, v^c \in V_{II} \wedge v^c \notin V_{II}^a\}$ ;  $V^c \leftarrow V_{II}^c$ ;
8       $V^v \leftarrow \{v | v \rightarrow T, \forall T \in V_{II}^a\}$ ;
9      else  $V^v \leftarrow \{v | v \rightarrow T^c\}$  for  $\forall v \in V^v$  do
10       if  $\text{FILTER} \in A(v)$  then
11         Arranges all  $\text{FILTER}$  conditions;
12       if  $\text{OPT} \in A(v)$  then
13         if  $(\text{LCA}(v) = \text{OPT}) \vee (V^c \subseteq V_{II}^a)$  then
14           break;
15         else continue
16       Chosen  $T^{(i)}$  from  $V^c$  using H1, H2;
17        $V_{II}^a \leftarrow T^{(i)}$ ;  $T^c \leftarrow T^{(i)}$ ;
18       if  $T^c \in R(\odot) \wedge T^{(i)} \in R(\odot)$  then
19         Appends Region operator  $\odot$  and  $T^{(i)}$  to  $\mathcal{CS}$ ;
20       else
21         Checks LCA operator, and adds the operators in
           the paths from LCA to  $T^c$  and  $T^{(i)}$  to  $\mathcal{CS}$ ;

```

a bit complicated. It is achieved by checking all connected variables in V^v . If all of these variables are *available*, or the LCA of these variables are OPT , then all the connected TPs can be arranged using H4. (Line 12-15).

In the *Generation* phase, a \mathcal{CS} is incrementally generated by appending the chosen TP $T^{(i)}$ and its related operators to the preceeding \mathcal{CS} . If $T^{(i)}$ and T^c are in the same *Region* $R(\odot)$, they can be directly added with the operator \odot using H1 and H2 (Line 19). If they are from different *Regions*, there a traversal of the operator vertex $v \in V_I$ is needed for LCA determination. All the traversed vertices are arranged in \mathcal{CS} accordingly (Line 21). It is noteworthy that such LCA-based method is also extensively adopted in traditional query plan generation. The process of Line 3-21 iterates until all TPs and operators have been arranged in a \mathcal{CS} .

EXAMPLE 5. Consider the QRG in Figure 3. T_5 is first chosen according to H1. Then the $R(\text{AND}_2)$ that contains T_5 is explored. As T_4 and T_6 are all connect to vertex $?p1$ with the same edge label S , they are chosen accordingly using H2. There T_6 also connects to vertex $?pc$, and $?pc$ has a FILTER connected. According to H3, compared with T_3 , C and FILTER is arranged with priority. So far, we have got a partial $\mathcal{CS} = (((T_1 \text{ AND } T_4) \text{ AND } (T_6 \text{ FILTER } C)) \text{ AND } T_3)$, with $?p1, ?pc$ and $?u1$ available. Next, T_1 and T_7 in $R(\text{OR})$ is considered. There according to H4, $R(\text{AND}_4)$ will not be considered until $?p1, ?u1$ and $?u2$ are available in all *Regions*. So T_{11} and T_{10} are arranged lately. Incrementally, we get a final $\mathcal{CS} = ((((((T_5 \text{ AND } T_4) \text{ AND } (T_6 \text{ FILTER } C)) \text{ AND } T_3) \text{ AND } (T_7 \text{ OR } (T_9 \text{ AND } T_8))) \text{ OPT } (T_{11} \text{ AND } T_{10})))$.

5.3 \mathcal{CS} -to-SQL Conversion

Intuitively, a \mathcal{CS} represents an operator tree, with TPs as tree leafs, and operators as internal nodes. As stated in [13], the semantics of these operators are analogous between SPARQL and SQL. Specifically, there are direct mapping of OR to an algebraic *Union*, OPTIONAL to a LEFT OUTER JOIN , AND to a *Join*, and FILTER with constraints

can be regarded as a *Selection*. Supporting comprehensive SPARQL specification is a non-trivial task. This needs semantics-preserving SPARQL to relational algebra conversion, and is beyond the scope of this paper. There we focus on the SELECT query with AND , OR , OPTIONAL and FILTER operators. Such simplifications are pragmatic and can cover substantially all real-world SPARQL queries, as stated in [31]. Besides, there are equivalent specifications in both SPARQL and SQL. i.e. ORDER BY , DISTINCT , LIMIT , OFFSET , NOT EXISTS , and MINUS . All these can be natively supported by ROSIE.

6. RUNTIME OPTIMIZATION

Due to the lack of precise statistics, the optimality of a derived \mathcal{CS} can not be guaranteed. The effectiveness and efficacy of such runtime optimization relies on the discrimination of unacceptable biased steps in a \mathcal{CS} . In this section, we introduce an optimization technique based on approximate error propagation.

6.1 Basic ideas

Generally, based on the definition of cardinality estimation error (Equation 2), the following *Properties* describe the propagation of errors for different operators.

- PROPERTY 3.** The cardinality error can be estimated as,
1. $\varepsilon(\mathcal{CS}_j \text{ AND } \mathcal{CS}_k) = \varepsilon(\mathcal{CS}_j) \times \varepsilon(\mathcal{CS}_k) \times \varepsilon(p(\mathcal{CS}_j \text{ AND } \mathcal{CS}_k))$;
 2. $\varepsilon(\mathcal{CS}_j \text{ OPT } \mathcal{CS}_k) = \varepsilon(\mathcal{CS}_j) \times \varepsilon(\mathcal{CS}_k) \times \varepsilon(p(\mathcal{CS}_j \text{ OPT } \mathcal{CS}_k))$;
 3. $\varepsilon(\mathcal{CS}_j \text{ OR } \mathcal{CS}_k) \leq \max(\varepsilon(\mathcal{CS}_j), \varepsilon(\mathcal{CS}_k))$;
 4. $\varepsilon(\mathcal{CS}_j \text{ FILTER } C) = \varepsilon(p(C))$.

where \mathcal{CS}_i and \mathcal{CS}_j are two partial \mathcal{CS} , and C is a condition expression for FILTER .

To better illustrate our ideas, we consider a fundamental circumstance that only $\odot = \text{AND}$ involved. Lets \mathcal{CS}_i denotes a \mathcal{CS} at step i , and $\widehat{\mathcal{CS}}_i$ is one of its equivalence. We consider three consecutive steps \mathcal{CS}_{i-1} , \mathcal{CS}_i and \mathcal{CS}_{i+1} that have:

$$\begin{aligned} \mathcal{CS}_{i+1} &= \mathcal{CS}_i \odot T_{i+1} = ((\mathcal{CS}_{i-1} \odot T_i) \odot T_{i+1}); \\ \widehat{\mathcal{CS}}_{i+1} &= \widehat{\mathcal{CS}}_i \odot T_i = ((\mathcal{CS}_{i-1} \odot T_{i+1}) \odot T_i). \end{aligned}$$

where T_i and T_{i+1} are TPs. As \mathcal{CS}_{i+1} is assumed to be optimal, then the condition $|\mathcal{CS}_{i+1}|_E \leq |\widehat{\mathcal{CS}}_{i+1}|_E$ holds. As \odot satisfies Property 1, then $\mathcal{CS}_{i+1} = \widehat{\mathcal{CS}}_{i+1}$, and we have $|\mathcal{CS}_{i+1}|_P = |\widehat{\mathcal{CS}}_{i+1}|_P$. Lets ε_{i+1} and $\widehat{\varepsilon}_{i+1}$ denote the estimation error of $|\mathcal{CS}_{i+1}|$ and $|\widehat{\mathcal{CS}}_{i+1}|$ respectively, we get:

$$\frac{|\mathcal{CS}_{i+1}|_P}{|\widehat{\mathcal{CS}}_{i+1}|_E} \leq \frac{|\widehat{\mathcal{CS}}_{i+1}|_P}{|\widehat{\mathcal{CS}}_{i+1}|_E} \Rightarrow \varepsilon_{i+1} \leq \widehat{\varepsilon}_{i+1} \quad (3)$$

Equation 3 states that the error is not propagated in a relatively incremental fashion in an optimal \mathcal{CS} . If Equation 3 satisfied, we can say that an optimal \mathcal{CS} is greedily chosen at step i . Otherwise, we need to re-consider the next i th step by materializing \mathcal{CS}_{i-1} . Though the idea is straightforward, actually the error ε and $\widehat{\varepsilon}$ at step i and $i+1$ are not readily available, as the actual cardinality $|\cdot|_P$ is unknown. In fact, derive the exact cardinalities are costly and in most cases unrealistic. But rely on the available statistics, a theoretical bound of $|\cdot|_P$ can be estimated, which is useful in determining the possible error-prone step in runtime optimization. Next, we explain the method for error bounds estimation by means of $|\cdot|_P$ bounds determination.

6.2 Error Bound Estimation

Basically, overestimating the error leads to unnecessary materialization steps, while underestimating the error can lead to suboptimal \mathcal{CS} . As materialization is a costly operation, and independence assumption tends to underestimate results, it is reasonable to only consider the case of underestimating. As discussed before, in a RDB, the precise cardinality of each distinct elements in \mathcal{D} is generally precomputed and maintained as single-value histogram. We denote these statistics as $|s|$, $|o|$ and $|p|$ for S, P and O respectively. The problems turns to, given $|s|$, $|o|$ and $|p|$, how to estimate a $|\mathcal{CS}|_P$.

We start with the cardinality estimate error of a single TP. There are seven types of TPs, i.e., $(s \ p \ ?o)$, $(?s \ p \ o)$, $(s \ ?p \ o)$, $(?s \ p \ ?o)$, $(s \ ?p \ ?o)$, $(?s \ ?p \ o)$ and $(?s \ ?p \ ?o)$. Intuitively, the cardinality of a TP that contains more than one unbound variables can be exactly captured. i.e. $|p|$, $|s|$, $|o|$ and $|\mathcal{D}|$ for $(?s \ p \ ?o)$, $(s \ ?p \ ?o)$, $(?s \ ?p \ o)$ and $(?s \ ?p \ ?o)$ respectively. It is a bit complicated for patterns that contain only one unbound variable, as S, P, O can not be assumed as mutually independent. It seems impossible to pre-compute correlations in full generality with reasonable efforts. We resort to estimate a bound for the selectivity, as the correlation varies from a functional dependency to complete independence. To be specific, taken pattern $T \in (?s \ p \ o)$ for example, $|T|_P = p \cdot (p = p) \cdot p \cdot (o = o \mid p = p) \cdot |\mathcal{D}| \in [\max(1, \frac{|p| \cdot |o|}{|\mathcal{D}|}), \min(|p|, |o|)]$, where the tighter lower bound comes from the independence assumption, and the upper bound states a functional independence using containment assumption⁸. This leads to a slight over-estimation, but in practice, it is far more accurate than assuming independence. Similarly, we can get the bound as follows:

$$|T|_P \in \begin{cases} [\max(1, \frac{|p| \cdot |s|}{|\mathcal{D}|}), \min(|p|, |s|)], & T \in (s \ p \ ?o); \\ [\max(1, \frac{|p| \cdot |o|}{|\mathcal{D}|}), \min(|p|, |o|)], & T \in (?s \ p \ o); \\ [\max(1, \frac{|s| \cdot |o|}{|\mathcal{D}|}), \min(|s|, |o|)], & T \in (s \ ?p \ o); \end{cases} \quad (4)$$

Next, we introduce operators to TPs. To better explain, we focus on JOIN, as it shows a multiplicative nature in error propagation. An accurate cardinality estimate of JOINS are difficult, as gathering and maintaining the statistics of \mathcal{D} on all possible joinable TPs is an expensive task, even if optimizations such as introducing the semantics of data are considered. Luckily, compared with cardinality estimate, error estimate is a more moderate problem, as it does not needs high accuracy as cardinality estimate does in CBO. Thus it is feasible to derive a reasonable error bound with lower cost. To get a tighter bound, we try to exploit as much information about correlations as possible. Consider the partial $\widetilde{\mathcal{CS}} = ((T_m \text{ AND } \dots) \text{ AND } T_n)$. The cardinality of $\widetilde{\mathcal{CS}}$ can be estimated as:

$$|\widetilde{\mathcal{CS}}|_P = (\prod_{i=m}^{n-1} p(T_i \text{ AND } T_{i+1})) \cdot (\prod_{i=m}^n |T_i|) \quad (5)$$

where $p(T_i \text{ AND } T_{i+1})$ is the selectivity of the JOIN predicate between T_i and T_{i+1} , if it exists. Besides seven type of TPs, there are six types of JOIN, namely SS-, SO-, OO-, SP-, OP- and PP-JOIN. Similarly, the lower and upper bound can be estimated under containment assumption and fully

correlated circumstance. That is,

$$p(T_i \text{ AND } T_{i+1}) \in \begin{cases} [\frac{1}{|T_i| \cdot |T_{i+1}|}, 1], & \text{AND is SO-/SP-/OP-JOIN;} \\ [\frac{1}{|T_i| \cdot |T_{i+1}|}, \frac{1}{\max(|T_i|, |T_{i+1}|)}], & \text{AND is SS-/OO-/PP-JOIN.} \end{cases} \quad (6)$$

The estimation of $\varepsilon(\widetilde{\mathcal{CS}})$ varies according to TP and JOIN types. Different combination of TP and JOIN types have different impact on error propagation. This should be considered individually using Equation 4, 6 and 5.

Similarity, we can estimate the error bound following Property 3, and revise \mathcal{CS} according to Property 2 or Property 1. They will not be detailed in this paper. In some cases this bound is not very useful, as the upper bound can be very large because it assumes fully correlation in data and derives a cardinality of Cartesian product. Without introducing extra statistical data which might change the histogram implementation of the underlying RDB, we will have to fall back to guessing at some point, and scale down the upper bound according to a predefined percentage. This has no theoretical guarantee but can greatly reduce the number of incremental evaluations.

EXAMPLE 6. For \mathcal{Q}_e in Figure 1, $\mathcal{CS} = ((T_4 \text{ AND } T_6) \text{ AND } T_3)$, and one of its equivalence $\widetilde{\mathcal{CS}} = ((T_4 \text{ AND } T_3) \text{ AND } T_6)$. The statistics held by RDB are $|\mathcal{D}| = 5000$, $|p = creator_of| = 200$, $|p = content| = 500$, $|p = type| = 300$ and $|o = Post| = 70$. The types of TP are $(?s \ p \ ?o)$ and $(?s \ p \ o)$, and the JOIN types includes SS- and SO-JOIN. According to Equation 4, $|T_4|_P \in [1, 70]$, $|T_3|_P = 200$, and $|T_6|_P = 500$. Applying Equation 6, we get $|\widetilde{\mathcal{CS}}|_P \in [1, 7 \times 10^6]$, while $|\mathcal{CS}|_P \in [1, 3.5 \times 10^4]$. As $|\widetilde{\mathcal{CS}}|_P = |\mathcal{CS}|_P$ held, we can infer that $\varepsilon(\mathcal{CS}) \geq \varepsilon(\widetilde{\mathcal{CS}})$. This states that \mathcal{CS} may not be not optimal.

7. EXPERIMENTS

In this section, we empirically evaluated the RDF query performance on two kinds of large-scale benchmark data, as well as real DBpedia dataset. The evaluations were focused on the quantitative metrics, i.e. the query response time.

7.1 Environments, Competitors and Datasets

Environments. All experiments were performed on a machine with Debian 7.4 in 64-bit Linux kernel, two Intel Xeon E5-2640 2.0GHz processors and 64GB RAM. ROSIE was developed in C++, and was compiled using GCC-4.7.1 with -O3 optimization flag. We used PostgreSQL 9.4 as RDB back-end, adopted the relational schema proposed by [7]. The default DB configuration file is deployed, only tuned the system parameter *shared_buffers* to a value of 2.5GB, which enabled more shared memory buffers used.

Datasets. We adopted two different kinds of benchmarks and real DBpedia dataset, as listed in Table 1.

Table 1: Statistics and characteristics of RDF dataset.

Dataset	LUBM	SNIB	DBpedia
# Triples	4,845M	1,774M	489M
# Predicates	18	48	60,223

LUBM⁹ is a widely used benchmark in both academical researches and industrial applications. It provides an university dataset where all components can be generated in a proportional fashion. We generated a large dataset of 20,480 universities. Remind that to support the original LUBM

⁸We implicitly assume that a TP T has a least one match, $|T| \geq 1$.

⁹<http://swat.cse.lehigh.edu/projects/lubm/>

Table 2: Canonical LUBM benchmark query response time (in milliseconds).

*: Wrong in query result size. N/A: Query does not ended in 12 hours. g.m. Geometric mean.

	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	g.m.
RDF-3X	2	2.3×10^6	3	5	7	5.8×10^6	7	401	830,944	2	362	14	38,988	3.8×10^6	718
TripleBit	0.2	56,005*	0.3	2	3	8.1×10^6	5	169*	527,032	0.5	3	4	639	7.7×10^6	147
Virtuoso7	2	42,521	1	3	4	6.3×10^6	4	402	440,644	3	3	24	3,759	4.5×10^6	278
DB2RDF	42	92,572	41	67	58	142,375	36	266	162,772	25	25	52	562	218,392	584
ROSIE-S	37	102,191	52	56	43	147,323	35	192	142,324	26	62	62	1,212	229,311	638
ROSIE	37	113,721	52	56	43	150,521	37	193	153,702	26	62	59	1,236	228,723	648
#Results	4	2,528	6	34	719	213,817,916	67	7,790	2,703,043	4	224	56	95,522	162,211,567	

Table 3: Response time of complex queries on LUBM and SNIB dataset (in milliseconds).

*: Wrong in query result size. N/A: Query does not ended in 12 hours. -: Query does not currently supported. g.m. Geometric mean.

	LUBM							SNIB					
	L15	L16	L17	L15 _{sel}	L16 _{sel}	L17 _{sel}	g.m.	S1	S2	S3	S4	S5	g.m.
RDF-3X	1.2 × 10 ⁶	N/A	N/A	3,783	14,968	92,775	50,107	2.4 × 10 ⁶	16,227	-	8,792	5.2 × 10 ⁶	205,416
TripleBit	1.0 × 10 ⁶ ★	456,658★	1.6 × 10 ⁶ ★	4,748★	5,890★	6,708	71,805	3,234	6,076	-	-	124,143★	13,416
Virtuoso7	1.4 × 10 ⁶	41,013	23,626	83	951	444	6,018	1,399	16,140	164,666	1,751	23,387	10,877
DB2RDF	192,662	76,391	32,442	283	562	342	5,441	3,527	12,213	57,326	1,207	11,812	8,115
ROSIE-S	522,079	452,394	627,048	3,202	4,253	6,342	48,359	8,423	120,832	512,089	23,258	85,482	63,546
ROSIE	84,323	29,603	17,920	127	472	383	3,176	8,712	4,232	21,306	1,411	6,923	5,907
#Results	17,200,845	1,116	21	720	10	3		27,318	54,062	258,707	323	21	

benchmark queries, the standard way is to inference on *subclasses* and *subproperties* schema. We used the inference engine in Virtuoso 7 to generated all inferred triples.

SNIB¹⁰ provides an RDF dataset of a Twitter-like social network that includes resources like users, post, reply, tags and comments, etc. The dataset scales according to the number of users. We generated a SNIB dataset of 15,000 users using S3G2 [30]. As it models a social network, SNIB data follows a power-law distribution.

DBpedia¹¹ contains information extracted from Wikipedia, and uses RDF as its data description format. Due to DBpedia data spreads on a broad spectrum of topics, its real-world queries are generally sophisticated in nature. There we used its English-only subsets.

Competitors. There exists a wide choice on the systems that support SPARQL. Among them, four publicly available and competitive ones were chosen. They can be classified as: *Dedicated Systems*: **RDF-3X**¹² maintains all possible permutation of *S*, *P* and *O* as indexes, and uses B+-tree to facilitate index lookup for single TP. With index-specific query optimization techniques, as well as runtime techniques like *sideways information passing*, RDF-3X remains a competitive state-of-the-art triple store [28]. **TripleBit**¹³ is based a novel data storage that use compact bit-matrices and vertical partitions. Its query optimizer employs a dynamical TP JOIN order that favors star-shaped sub-queries, and shows by far the best query performance compared with other dedicated systems [37].

RDB-backed Systems: **Virtuoso 7**¹⁴ is a commercial-of-the-shelf column store that supports RDF management. It exploits existing relational RDB techniques by adding functionalities to deal with SPARQL, and adopted a "vectorized query execution" technique as outlined in [14]. **DB2RDF**¹⁵ proposes an entity-oriented relational schema that of ex-

ploits the merits of property table but is more compact in physical storage. Based on that, it provides an effective SPARQL-to-SQL conversion method for RDF over RDB [7]. Compared with ROSIE, DB2RDF differs in that it focused on generating optimized SQL which is committed integrally, leaves the burden of all rest works to RDB.

7.2 Results Analysis

The queries used are listed in *Appendix*. Each query was executed 11 times in a consecutive manner in each competitor. As the first time of each round was used for cache warm-up, its results was removed from averaging. Thus we reported the query response time as the arithmetic mean of the rest 10 queries, and rounded them to milliseconds. For each experiment, we given the geometric mean of all queries' response times.

To better evaluate the effects of runtime optimization technique, we also implemented a simplified framework for comparison, denoted as **ROSIE-S**. ROSIE-S generated a static *CS* as described in *Section 5*, and used this throughout the query execution, without employed the mechanism described in *Section 6*.

(E-1) Canonical LUBM Benchmark Results. Table 2 showed the benchmark query performance on LUBM data. As all these benchmark query are relatively simple in that they contain fewer TPs and only AND operator involved, runtime optimization mechanism in ROSIE did not worked. Thus ROSIE and ROSIE-S showed approximately the same query performance. Query *L1*, *L3*, *L4*, *L5*, *L7*, *L8*, *L10*, *L11* exhibit the same characteristics in that they all have high-selective TPs in query. Besides, the result sets of these queries are rather small and are independent of the data size. Even though the RDB-backed methods, i.e., Virtuoso, DB2RDF and ROSIE, showed an inferior performance, consider that all the query response times are within 1 second, it is hard for a user to feel the difference. Query *L2*, *L6*, *L9*, *L13* and *L14* are characterized by returning a large size of results. The RDB-based methods performed well on these queries. This can be attributed to the efficiency of RDB in handling large amount of disk-based data.

(E-2) Complex Queries on Benchmark Data. Table 3 showed

¹⁰ http://www.w3.org/wiki/Social_Network_Intelligence_BenchMark

¹¹ Downloaded from <http://wiki.dbpedia.org/Downloads2015-04>

¹² Available at <https://github.com/gh-rdf3x/gh-rdf3x/>

¹³ Available at <http://grid.hust.edu.cn/triplebit/TripleBit.tar.gz>

¹⁴ Available at <https://github.com/openlink/virtuoso-opensource>

¹⁵ Available at <https://github.com/Quetzal-RDF/quetzal/>

the performance of more complex queries on LUBM and SNIB data. These queries are "complex" in that they typically contains more JOINS, range from 6 to 15, with OPT, UNION and FILTER operators involved. Specifically, Query *L15-L17* are JOIN-intensive queries, and ROSIE consistently outperformed others. Query *L15_{sel}-L17_{sel}* are derived by adding a high selective TP to the corresponding Query *L15-L17*. These high selective TPs gave explicit hints for the query optimizer, and are helpful in guiding the optimal query plan generation. ROSI generally outperformed on these queries. For SNIB benchmark, it generated a more correlated data than LUBM, and Query *S1-S5* on SNIB adopted more operators. Analogously, ROSIE achieved approximately the best performance. For all queries in E-2, ROSIE exhibited obvious superior query performances compared with ROISE-E. This demonstrated the effectiveness of the runtime optimization mechanism applied.

(E-3) Real Queries on DBpedia. Table 4 showed the performance of queries on DBpedia data. In case of real queries *D1-D6*, they are characterized by UNION-intensive (*D1*), Filter-intensive (*D3*) or Opt-intensive (*D6*). RDF-3X and TripleBit are more focused on AND queries, this renders them incapable for more expressive queries, and they were not employed in this experiment. Consider geometric means, ROSIE achieved approximately 1.5 times faster than DB2RDF, and 3 times faster than Virtuoso.

Table 4: Queries response time on DBpedia dataset.
(in milliseconds)

	D1	D2	D3	D4	D5	D6	g.m.
Virtuoso7	542	235	3,873	342	5,243	1,121	998
DB2RDF	72	742	5,702	92	2,734	462	573
ROSIE-S	121	2,253	7,311	812	6,723	1,889	1,655
ROSIE	62	212	3,102	84	1,078	372	332
#Results	1,182	11	58,761	53	48	2	

Results Analysis: To summarize, E-1 demonstrated the workability of ROSIE framework. The comparative performance of ROSIE with others showed the efficiency of the CS generated by the heuristic-based planning algorithm, and the results size showed the correctness of query execution in ROSIE. As shown in E-2 and E-3, for more complex queries on data with more sophisticated correlations, RDB-based systems outperformed the dedicated systems. ROSIE performed much better, achieved 1.5X to 20X better query performance. In addition, compared with ROSIE-S that used a static and heuristic-based query planning, the performance of ROSIE outperformed by orders of magnitudes. These shown that for complicated query and sophisticated data, the essence of ROSIE runtime optimization gets highlighted. We trust that with further optimizations applied for the back-end RDB, ROSIE still have room for boosting its query performance.

8. CONCLUSIONS

In this paper, we introduced ROSIE (Runtime Optimization of SPARQL query using Incremental Evaluation), a framework aiming at promoting the query performance, as well as supporting more expressive SPARQL queries. To achieve these, we proposed a heuristic-based approach for TP execution order determination. Furthermore, we devised a mechanism that adopted cardinality estimate error bound determination to optimize the generated order dynamically at run-

time. We showed the correctness and effectiveness of ROSIE through extensive experiments on synthetic and real RDF data. There are still many open issues left for future work. Among them, supporting queries that involved more complete SPARQL specifications, and the semantic-preserving SPARQL-to-SQL conversion are the most promising aspects to be explored.

Acknowledgments

<Intentionally Left Blank>

9. REFERENCES

- [1] DBpedia. <http://wiki.dbpedia.org/>
- [2] Freebase. <https://www.freebase.com/>
- [3] UniProt. <ftp://ftp.uniprot.org/>
- [4] Abadi, D.J., Marcus, A., Madden, S.R., Hollenbach, K.: Scalable semantic web data management using vertical partitioning. In: VLDB. pp. 411–422 (2007)
- [5] Abadi, D.J., Marcus, A., Madden, S.R., Hollenbach, K.: SW-Store: A vertically partitioned DBMS for semantic web data management. PVLDB 18(2), 385–406 (2009)
- [6] Abdel Kader, R., Boncz, P., Manegold, S., van Keulen, M.: ROX: Run-time optimization of XQueries. In: SIGMOD. pp. 615–626 (2009)
- [7] Bornea, M., Dolby, J., Kementsietsidis, A., Srinivas, K., Dantressangle, P., Udrea, O., Bhattacharjee, B.: Building an efficient RDF store over a relational database. In: SIGMOD. pp. 121–132 (2013)
- [8] Broekstra, J., Kampman, A., Harmelen, F.v.: Sesame: A generic architecture for storing and querying RDF and RDF schema. In: ISWC. pp. 54–68 (2002)
- [9] Bruno, N., Chaudhuri, S.: Exploiting statistics on query expressions for optimization. In: SIGMOD. pp. 263–274 (2002)
- [10] Chaudhuri, S.: An overview of query optimization in relational systems. In: PODS. pp. 34–43 (1998)
- [11] Chong, E.L., Das, S., Eadon, G., Srinivasan, J.: An efficient SQL-based RDF querying scheme. In: VLDB. pp. 1216–1227 (2005)
- [12] Cole, R.L., Graefe, G.: Optimization of dynamic query evaluation plans. In: SIGMOD. pp. 150–160 (1994)
- [13] Eric, P., Alexandre, B.: A mapping of SPARQL onto conventional SQL. <https://www.w3.org/2008/07/MappingRules/StemMapping> (2008)
- [14] Erling, O., Mikhailov, I.: Virtuoso: RDF support in a native RDBMS. In: Semantic Web Information Management, pp. 501–519. Springer (2009)
- [15] Graefe, G., Ward, K.: Dynamic query evaluation plans. In: SIGMOD. pp. 358–366 (1989)
- [16] Gubichev, A., Neumann, T.: Exploiting the query structure for efficient join ordering in SPARQL queries. In: EDBT. pp. 439–450 (2014)
- [17] Harris, S., Seaborne, A.: SPARQL 1.1 query language. W3C working draft. <http://www.w3.org/TR/sparql11-query> (2013)
- [18] Harth, A., Hose, K., Schenkel, R.: Database techniques for linked data management. In: SIGMOD. pp. 597–600 (2012)

- [19] Ioannidis, Y.E.: The history of histograms (abridged). In: VLDB. pp. 19–30 (2003)
- [20] Ioannidis, Y.E., Christodoulakis, S.: On the propagation of errors in the size of join results. In: SIGMOD. pp. 268–277 (1991)
- [21] Ioannidis, Y.E., Christodoulakis, S.: Optimal histograms for limiting worst-case error propagation in the size of join results. ACM Trans. Database Syst. 18(4), 709–748 (Dec 1993)
- [22] Kotoulas, S., Urbani, J., Boncz, P., Mika, P.: Robust runtime optimization and skew-resistant execution of analytical sparql queries on pig. In: ISWC. pp. 247–262 (2012)
- [23] Leis, V., Gubichev, A., Mirchev, A., Boncz, P., Kemper, A., Neumann, T.: How good are query optimizers, really? PVLDB 9(3) (Nov 2015)
- [24] Lipton, R.J., Naughton, J.F., Schneider, D.A.: Practical selectivity estimation through adaptive sampling. In: SIGMOD. pp. 1–11 (1990)
- [25] Lohman, G.: Is query optimization a solved problem? <http://wp.sigmod.org/?p=1075> (2014)
- [26] Neumann, T., Galindo-Legaria, C.A.: Taking the edge off cardinality estimation errors using incremental execution. In: BTW. pp. 73–92 (2013)
- [27] Neumann, T., Moerkotte, G.: Characteristic sets: Accurate cardinality estimation for RDF queries with multiple joins. In: ICDE. pp. 984–994 (2011)
- [28] Neumann, T., Weikum, G.: The RDF-3X engine for scalable management of RDF data. In: PVLDB. vol. 19, pp. 91–113 (2010)
- [29] Pérez, J., Arenas, M., Gutierrez, C.: Semantics and complexity of SPARQL. ACM Trans. Database Syst. 34(3), 16:1–16:45 (2009)
- [30] Pham, M., Boncz, P.A., Erling, O.: S3G2: A scalable structure-correlated social graph generator. In: TPCTC. pp. 156–172 (2012)
- [31] Picalausa, F., Vansummeren, S.: What are real SPARQL queries like? In: SWIM. pp. 7:1–7:6 (2011)
- [32] Schmidt, M., Meier, M., Lausen, G.: Foundations of SPARQL query optimization. In: ICDT. pp. 4–33 (2010)
- [33] Stocker, M., Seaborne, A., Bernstein, A., Kiefer, C., Reynolds, D.: SPARQL basic graph pattern optimization using selectivity estimation. In: WWW. pp. 595–604 (2008)
- [34] Tsialiamanis, P., Sidiropoulos, L., Fundulaki, I., Christophides, V., Boncz, P.: Heuristics-based query optimisation for SPARQL. In: EDBT. pp. 324–335 (2012)
- [35] Weiss, C., Karras, P., Bernstein, A.: Hexastore: Sextuple indexing for semantic web data management. PVLDB. 1(1) (Aug 2008)
- [36] Wilkinson, K.: Jena Property Table Implementation. In: SSWS. pp. 35–46 (2006)
- [37] Yuan, P., Liu, P., Wu, B., Jin, H., Zhang, W., Liu, L.: TripleBit: A fast and compact system for large scale RDF data. PVLDB 6(7), 517–528 (2013)

10. APPENDIX

10.1 LUBM Queries

PREFIX rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>
 PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>
 PREFIX ub: <<http://swat.cse.lehigh.edu/onto/univ-bench.owl#>>
L1-L14: Same as LUBM Q1-Q14 at <http://swat.cse.lehigh.edu/projects/lubm/queries> respectively.
L15: SELECT ?a1 ?a2 ?a3 ?a4 WHERE { ?a4 rdf:type ub:University. ?a1 rdf:type ub:GraduateStudent. ?a1 ub:advisor ?a2. ?a2 ub:worksFor ?a3. ?a3 ub:subOrganizationOf ?a4. ?a2 rdf:type ub:FullProfessor. ?a3 rdf:type ub:Department. }
L15_{sel}: L15 with pattern { ?a4 ub:name "University7" }.
L16: SELECT ?x ?y ?z WHERE { ?x rdf:type ub:GraduateStudent. ?y rdf:type ub:University. ?z rdf:type ub:Department. ?c rdf:type ub:GraduateCourse. ?x ub:takesCourse ?c. ?p rdf:type ub:FullProfessor. ?p ub:teacherOf ?c. ?p ub:worksFor ?z. ?x ub:memberOf ?z. ?x ub:undergraduateDegreeFrom ?y. ?z ub:subOrganizationOf ?y }
L16_{sel}: L16 with pattern { ?y ub:name "University6" }.
L17: SELECT ?x ?y ?p ?b ?c WHERE { ?x rdf:type ub:GraduateStudent. ?x rdf:type ub:TeachingAssistant. ?x ub:advisor ?p. ?x ub:takesCourse ?c. ?x ub:memberOf ?z. ?x ub:undergraduateDegreeFrom ?y. ?y rdf:type ub:University. ?c rdf:type ub:GraduateCourse. ?p ub:worksFor ?z. ?p rdf:type ub:AssociateProfessor. ?z ub:subOrganizationOf ?y. ?z rdf:type ub:Department. ?b rdf:type ub:Publication. ?p ub:teacherOf ?c. ?b ub:publicationAuthor ?x. }
L17_{sel}: L17 with pattern { ?Y ub:name "University786" }.

10.2 SNIB Queries

PREFIX foaf: <<http://xmlns.com/foaf/0.1/>>
 PREFIX dc: <<http://purl.org/dc/elements/1.1/>>
 PREFIX sioc: <<http://rdfs.org/sioc/ns#>>
 PREFIX sioc: <<http://rdfs.org/sioc/type#>>
 PREFIX sib: <<http://www.ins.cwi.nl/sib/vocabulary/>>
 PREFIX dbp: <<http://dbpedia.org/resource/>>
S1: SELECT ?user ?commentcontent ?commentdate WHERE { ?user foaf:knows ?friend. ?user rdf:type sib:User. ?friend rdf:type sib:User. ?friend sioc:moderator_of ?forum. ?post sib:hashtag dbp:Creek. ?post rdf:type sioc:Post. ?forum sioc:container_of ?post. ?post sioc:content ?postcontent. ?post sioc:container_of ?postcomment. ?postcomment rdf:type sioc:Item. ?postcomment sioc:content ?commentcontent. ?postcomment dc:created ?commentdate. }
S2: SELECT ?user1 ?user2 ?friend WHERE { ?user1 foaf:knows ?user2. ?user2 foaf:knows ?friend. ?friend foaf:knows ?user1. ?user1 rdf:type sib:User. ?user2 rdf:type sib:User. ?friend rdf:type sib:User. ?friend sioc:creator_of ?post. ?post rdf:type sioc:Post. ?post sioc:content ?postcontent. ?post sib:hashtag dbp:Creek. ?post sib:liked_by ?user1 }
S3: SELECT ?user0 ?friend ?photo WHERE { ?user0 rdfs:type sib:User. ?friend rdfs:type sib:User. ?photo rdfs:type sib:Photo. { ?friend foaf:knows ?user0. } OPTIONAL { ?friend foaf:knows ?u2. ?u2 foaf:knows ?user0. } ?photo sib:usertag ?friend. ?photo dbp:location "Germany". ?user0 sib:liked_by ?photo. FILTER EXISTS { ?pa sioc:container_of ?photo. ?pa rdfs:type type:ImageGallery. { ?u2 sioc:creator_of ?pa. } UNION { ?user0 sioc:creator_of ?pa. } } }
S4: SELECT ?user1 WHERE { ?user1 rdf:type sib:User. ?user2 rdf:type sib:User. ?photo rdfs:type sib:Photo. ?user2 sioc:creator_of ?pa. ?photo sib:usertag ?user1. ?pa rdfs:type sioc:ImageGallery. ?user1 sib:InRelationshipWith ?user2. ?pa sioc:container_of ?photo. ?user2 sioc:creator_of ?pa. OPTIONAL { ?post sib:liked_by ?user1. ?post sioc:content ?postcontent. ?post sib:hashtag dbp:Creek. } }
S5: SELECT ?u1 ?u2 WHERE { ?f sib:memb ?u1. ?f sib:memb ?u2. ?u2 foaf:knows ?u1. ?u3 foaf:knows ?u2. ?u3 foaf:knows ?u1. ?u1 rdf:type sib:User. ?u2 rdf:type sib:User. ?u3 rdf:type sib:User. ?p sib:hashtag dbp:Hummelshof. ?u1 sioc:creator_of ?p. ?p rdf:type sioc:Post. ?forum sioc:container_of ?p. ?forum sioc:container_of ?c. ?c rdf:type sioc:Item. ?u2 sioc:creator_of ?c. ?c sioc:reply_of ?p. ?p sib:liked_by ?u2. }

10.3 DBpedia Queries

PREFIX dbp: <<http://dbpedia.org/>>
 PREFIX r: <<http://dbpedia.org/resource/>>
 PREFIX p: <<http://dbpedia.org/property/>>
 PREFIX o: <<http://dbpedia.org/ontology/>>
 PREFIX owl: <<http://www.w3.org/2002/07/owl#>>
 PREFIX xsd: <<http://www.w3.org/2001/XMLSchema#>>
 PREFIX skos: <<http://www.w3.org/2004/02/skos/core#>>
D1: SELECT ?club WHERE { { ?club p:league r:Premier_League. } UNION { ?club p:league r:Scottish_Premier_League. } UNION { ?club p:league r:Football_League_One. } UNION { ?club p:league r:Football_League_Championship. } UNION { ?club p:league r:Football_League_Two. } }
D2: SELECT DISTINCT ?v1 ?v2 WHERE { ?v1 rdf:type o:Wrestler. ?v2 rdfs:label ?v2. FILTER regex(str(?v2), "sep", "i") }
D3: SELECT ?v1 ?v2 WHERE { { ?v1 rdf:type o:Settlement. ?v1

```

p:population ?v2. FILTER ( xsd:integer(?v2) > 54 ) } UNION {
?v1 rdf:type o:Settlement. ?v1 p:populationUrban ?v2. FILTER (
xsd:integer(?v2) > 54 ) } }
D4:SELECT ?v1 ?v2 ?v3 ?v4 ?v5 WHERE { { r:Treehouse_of_Horror
_XX ?v1 ?v2. ?v2 foaf:name ?v4. } UNION { ?v3 ?v1 r:Treehouse_of
_Horror_XX ; ?v3 foaf:name ?v5. } }
D5:SELECT ?v1 ?v2 ?v3 ?v4 ?v5 WHERE{ ?v1 rdf:type o:Artist.
?v1 p:name ?v2. ?v1 p:pages ?v3. ?v1 p:isbn ?v4. ?v1 p:author ?v5.}
D6:SELECT ?V1 ?v2 ?v3 WHERE { ?v1 rdf:type o:Settlement. ?v1
rdfs:label "Djanet". ?v2 rdf:type o:Airport. { ?v2 o:city ?v1.} UNION
{ ?v2 o:iataLocationIdentifier ?v3. } UNION { ?v2 o:location ?v1. }
{ ?v2 p:iata ?v3. } OPTIONAL { ?v2 foaf:homepage ?v5. } OP-
TIONAL { ?v2 p:nativename ?v4. } }

```